DOCUMENT RESUME

ED 173 387 TM 009 245

AUTHOR Wilson, Pamela W.; And Others
TITLE An Empirical Adjustment of the Kuder-Richardson 21 Reliability Coefficient to Better Estimate the Kuder-Richardson 20 Coefficient.
INSTITUTION Michigan State Univ., East Lansing.
PUB DATE [Apr 79]
NOTE 9p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Higher Education; Item Analysis; *Mathematical Formulas; *Test Reliability
IDENTIFIERS *Kuder Richardson Reliability Formulas

ABSTRACT

The purpose of this study was to present an empirical correction of the KR21 (Kuder Richardson test reliability) formula that not only yields a closer approximation to the numerical value of the KR20 without overestimation, but also simplifies computation. This correction was accomplished by introducing several correction factors to the numerator of the KR21 to reduce the inflation of the error-variance estimate. The reliability coefficients resulting from this analysis were compared under varying levels of student sample size, number of items in the test, examination mean difficulty, and examination mean discrimination. Test data were taken from a representative sample of 101 university-level classroom achievement tests. The hypothesis of no difference between the mean KR20 and KR21 coefficients was rejected, while the same hypothesis for the empirically adjusted KR21 and the KR20 was retained. The mean underestimate of the KR20 by the KR21 was 16.1 percent, while the corrected formula 21 underestimated by only 4.2 percent. The .8 correction factor was superior to all other factors and to the original KR21, as estimates of the KR20. The representativeness of the sample of tests used in this study suggested that these findings are generalizable. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

An Empirical Adjustment of the Kuder-Richardson 21 Reliability Coefficient to Better Estimate the Kuder-Richardson 20 Coefficient

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Pamela W. Wilson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Abstract

Pamela W. Wilson
Steven M. Downing
Robert L. Ebel
Michigan State University

The goal of this study was to empirically correct the KR21 formula to more closely approximate the numerical value of formula 20. An n = 101 random sample of summary objective-achievement test data were selected. The hypothesis of no difference between the mean KR20 and KR21 coefficients was rejected ($p < .10$), while the same hypothesis for the empirically adjusted KR21 and the KR20 was retained. The mean underestimate of the KR20 by the KR21 was 16.1 percent, while the corrected formula 21 underestimated by only 4.2 percent. Other results are presented and the KR21 and its correction are plotted against the KR20.

ED173387

TM009 245

An Empirical Adjustment of the Kuder-Richardson 21 Reliability Coefficient to Better Estimate the Kuder-Richardson 20 Coefficient

Objectives of the Study

It is well known that the Kuder-Richardson Formula 21 (KR21) estimate of test reliability almost always underestimates the reliability coefficient given by the Kuder-Richardson Formula 20 (KR20). The numerical difference between these two formulas arises from the rigid assumption of equal item difficulties required for the KR21 derivation. The KR21's underestimation of the KR20 coefficient stems from violation of the assumption of equal item p-values. (Kuder and Richardson, 1937)

The usefulness of the KR21 formula arises from its ease of calculation (compared to the calculation of the KR20 formula). Since the KR21 requires only the test mean, variance, and number of test items for calculation, it is a coefficient that a classroom teacher finds useful. However, in using the KR21 the classroom teacher must be prepared to trade the accuracy of the KR20 for the computational ease of the KR21.

The purpose of this study is to present an empirical correction of the KR21 formula that not only yields a closer approximation to the numerical value of the KR20 without overestimation, but also simplifies computation. This correction will be accomplished by introducing a constant in the numerator of the KR21 to reduce the inflation of the error-variance estimate.

The proposed adjustment of the KR21 formula ignores the fraction $\frac{K}{K-1}$ and inserts a constant of .8 in the numerator.

$$\text{Adjusted KR21} = 1 - \frac{.8(\bar{X})\ (K - \bar{X})}{K(S^2_x)} \qquad (1)$$

where: $\bar{X}$ = test mean
K_2 = number of test items
S^2_x = test variance

Although the Kuder-Richardson formulas have been the subject of much theorizing and some modification (Horst, 1953; Loevinger, 1947; Lenke et al, 1977), no one has attempted to find a simple correction of the KR21 to better estimate the KR20 formula.

Data Source

The data for this study consist of a random sample of n = 101 examinations processed by a scoring office of a large midwestern university. Data were transcribed from a summary item analysis which included the KR20 reliability coefficient, the test mean, variance, number of test items and mean item difficulty and discrimination. All test data were from objective-format, university-level classroom achievement tests.

The exams varied in many respects. The length of the exams ranged from 13 to 132 items, the number of students taking each test ranged from 14 to 229, the mean item difficulty (mean proportion incorrect) of these tests ranged from 9 to 50, the mean discrimination index (D) of these exams ranged from 14 to 47, and the KR20 reliability coefficients ranged from .1355 to .9366.

Methodology

A Fortran program was written to compute the KR21 and a variety of adjusted KR21 formulas for each exam in the sample. The basic formula is presented below.

$$\text{Adjusted KR21} = 1 - \frac{C(\bar{X})\ (K - \bar{X})}{K\ (S^2_x)} \qquad (2)$$

where: C = correction factor

Several correction factors were investigated. They were C = .60, .65, .70, .75, .80, .85, .90, .95. The basic goal of this analysis was to find a constant that would improve the KR21 as an estimate of the KR20 without overestimating it.

The coefficients resulting from this analysis were then compared to determine a decision rule for applying various correction factors to the KR21 formula in different circumstances. The coefficients were compared under varying levels of student sample size (0-25, 26-50, 51-100, greater than 100), number of items in the test (less than 50-item test, greater than 50-item test), examination mean difficulty (less than 20, 21-40, 41-60, greater than 60), and examination mean discrimination (less than 20, 21-40, 41-60, greater than 60).

Two hypotheses were tested in this study:

1. H_0: The mean KR21 value is equal to the mean KR20 value.
   H_1: The mean KR21 value is less than the mean KR20 value.
2. H_0: The mean KR21 value (adjusted by .8) is equal to the KR20 value.
   H_1: The mean KR21 value (adjusted by .8) is less than the KR20 value.

Results and Conclusions

Examination of the various potential correction factors revealed the .8 factor to be better than all other correction factors and the original KR21 as estimates of the KR20. Within rounding error, 92 percent of the .8 corrections of the KR21 were better than the original KR21 without overestimating the KR20 coefficient. On the average, the KR21 underestimated the KR20 coefficient by 16.1 percent; the KR21 adjusted by .8 (Formula 1) underestimated the KR20 by only 4.2 percent. Three percent of the KR21 coefficients had negative values; in each of these cases, Formula 1 gave a positive value for the reliability coefficient. In 99 percent of the cases, the .8 correction was better than the KR21 with only a slight overestimation in 7 percent of the cases. (The one test in which the KR21 was superior to the .8 adjustment was in the case of a 10-item test; the KR21 was better by only .003.) Most (87 percent) of the overestimates occurred in shorter

exams (fewer than 50 items). However, there appeared to be no differences among the distribution of best estimates with regard to size of sample, mean item difficulty, or mean item discrimination. This finding supports Lenke et al (1977) which showed that the difference between the KR20 and KR21 is significantly greater when the item difficulty levels are less than .30.

Table 1

Comparisons of the KR20 with the KR21
and the KR20 with the Adjusted KR21
for n = 101 tests

	Mean	S.D.	Correlation	t-value*
KR20	.7170	0.150		
KR21	.6276	0.217	.930	-1.597**
KR20	.7170	0.150		
KR21 (Adjusted by .8)	.6912	0.173	.935	-0.509 NS

* Fisher's r to Z transformation
** $p < .10$

Table 1 shows that the first hypothesis of no difference between the means of the KR20 and KR21 was rejected at an alpha level of .10. The second hypothesis of no difference between the mean KR20 and the mean corrected KR21 (.8) was not rejected (at $\alpha = .10$).

There was, on the average, .09 units difference between the KR20 and

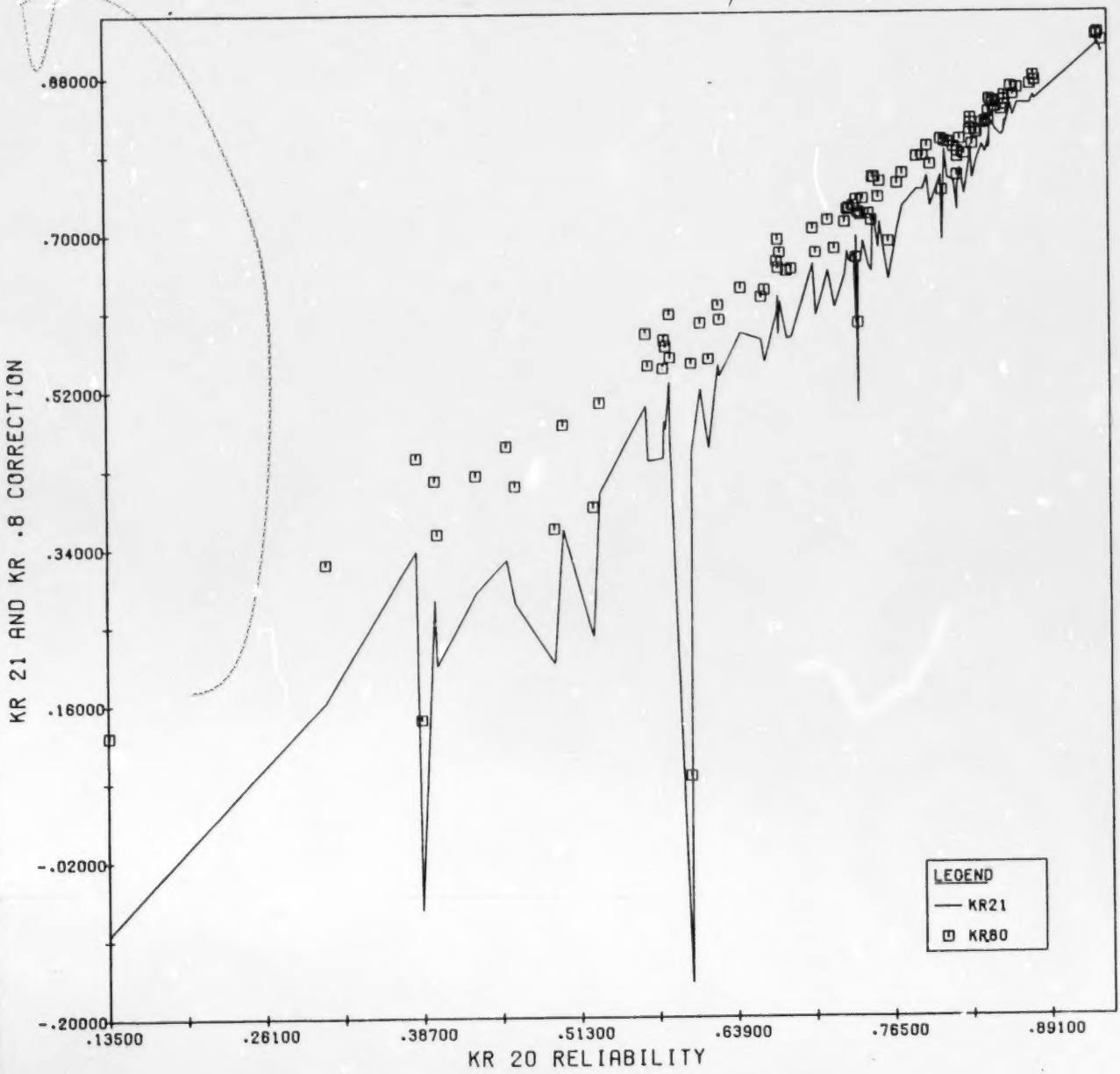
KR21, while the KR20 and the corrected KR21 differed by only .03 units. Additionally, the .8 corrected KR21 was slightly more highly correlated ($r = .935$) with the KR20 than was the KR21 with the KR20 ($r = .930$).

These results suggest that the adjustment of the KR21 formula (given by Formula 1) is superior to the usual KR21 formula in its estimation of the KR20 coefficient.

The representativeness of the sample of objective-format achievement tests used in this study suggests that these findings are generalizable, and, therefore, useful in situations similar to those of this study.

Figure 1 plots the KR21 coefficients and Formula 1 against the KR20 reliabilities. This plot clearly shows the higher values for the adjusted KR21 (relative to the usual KR21 formula) and visually demonstrates the closer approximation of Formula 1 to the KR20.

FIG. 1: KR 21 AND CORRECTED (.8) KR 21 BY KR 20



Bibliography

Horst, P., "Correcting the Kuder-Richardson Reliability for Dispersion of Item Difficulties." Psychological Bulletin. 30: 371-374, 1953. Reprinted in Mehrens, W. A. and Ebel, R. L. (eds.), Principles of Educational and Psychological Measurement. Chicago: Rand McNally and Co., 1967.

Kuder, G. F., and Richardson, M. W., "The Theory of the Estimation of Test Reliability." Psychometrika. 2: 151-160, 1937. Reprinted in Mehrens, W. A. and Ebel, R. L. (eds.), Principles of Educational and Psychological Measurement. Chicago: Rand McNally and Co., 1967.

Lenke, J. M., et al, "Differences Between Kuder-Richardson Formula 20 and Formula 21 Reliability Coefficients for Short Tests with Different Item Variabilities." AERA Paper: New York, 1977.

Loevinger, J., "A Systematic Approach to the Construction and Evaluation of Test of Ability." Psychological Monograph. 64: no. 4, 1947.